IN THE SPECIFICATION:

Please replace the paragraph on page 11, lines 8-26, with the following rewritten paragraph:

A1

-- Query and individual sequences can be aligned using the methods and computer programs described above, and include BLAST 2.0 (National Center for Biotechnology Information, Bethesda, Maryland). See also Altschul, et al. *Nucleic Acids Res.* (1997) *25*:3389-3402. Another alignment algorithm is Fasta, available in the Genetics Computing Group (GCG) package, Madison, Wisconsin, USA, a wholly owned subsidiary of Oxford Molecular Group, Inc. Other techniques for alignment are described in Doolittle, *supra*. Preferably, an alignment program that permits gaps in the sequence is utilized to align the sequences. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See *Meth. Mol. Biol.* (1997) *70*: 173-187. Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. An alternative search strategy uses MPSRCH software, which runs on a MASPAR computer. MPSRCH uses a Smith-Waterman algorithm to score sequences on a massively parallel computer. This approach improves ability to identify sequences that are distantly related matches, and is especially tolerant of small gaps and nucleotide sequence errors. Amino acid sequences encoded by the provided polynucleotides can be used to search both protein and DNA databases. Incorporated herein by reference are all sequences that have been made public as of the filing date of this application by any of the DNA or protein sequence databases, including the patent databases (*e.g.*, GeneSeq). Also incorporated by reference are those sequences that have been submitted to these databases as of the filing date of the present application but not made public until after the filing date of the present application. --

Please replace the paragraph on page 14, lines 9-20, with the following rewritten paragraph:

A2

-- Profiles can be designed manually by (1) creating an MSA, which is an alignment of the amino acid sequence of members that belong to the family and (2) constructing a statistical representation of the alignment. Such methods are described, for example, in Birney *et al.*, *Nucl. Acid Res.* (1996) *24(14)*: 2730-2739. MSAs of some protein families and motifs are publicly available. For example, the Pfam database available from Washington University (St. Louis, Missouri) includes MSAs of 547 different families and motifs. These MSAs are described also in Sonnhammer *et al.*, *Proteins* (1997) *28*: 405-420. Other publicly available sources include those over

the world wide web provided by the European Molecular Biology Laboratory (Heidelberg, Germany). A brief description of these MSAs is reported in Pascarella *et al.*, *Prot. Eng.* (1996) *9(3)*:249-251. Techniques for building profiles from MSAs are described in Sonnhammer *et al.*, *supra*; Birney *et al.*, *supra*; and "Computer Methods for Macromolecular Sequence Analysis," *Methods in Enzymology* (1996) *266*, Doolittle, Academic Press, Inc., San Diego, California, USA. --

AZ cancl

Please replace the paragraph starting at page 24, line 18, with the following rewritten paragraph:

AB -- Mapping. Polynucleotides of the present invention can be used to identify a chromosome on which the corresponding gene resides. Such mapping can be useful in identifying the function of the polynucleotide-related gene by its proximity to other genes with known function. Function can also be assigned to the polynucleotide-related gene when particular syndromes or diseases map to the same chromosome. For example, use of polynucleotide probes in identification and quantification of nucleic acid sequence aberrations is described in USPN 5,783,387. An exemplary mapping method is fluorescence in situ hybridization (FISH), which facilitates comparative genomic hybridization to allow total genome assessment of changes in relative copy number of DNA sequences (see, e.g., Valdes *et al.*, *Methods in Molecular Biology* (1997) *68*:1). Polynucleotides can also be mapped to particular chromosomes using, for example, radiation hybrids or chromosome-specific hybrid panels. See Leach *et al.*, *Advances in Genetics*, (1995) *33*:63-99; Walter *et al.*, *Nature Genetics* (1994) *7*:22; Walter and Goodfellow, *Trends in Genetics* (1992) *9*:352. Panels for radiation hybrid mapping are available from Research Genetics, Inc., Huntsville, Alabama, USA. Databases for markers using various panels are publicly available via the world wide web from the Stanford Genome Center and The Whitehead Institute for Biomedical Research/MIT Center for Genome Research. The statistical program RHMAP can be used to construct a map based on the data from radiation hybridization with a measure of the relative likelihood of one order versus another. RHMAP is available from the University of Michigan, Center for Statistical Genetics, Ann Arbor, Michigan. In addition, commercial programs are available for identifying regions of chromosomes commonly associated with disease, such as cancer. --

Please replace the paragraph on page 46, lines 26-33, with the following rewritten paragraph:

A4

-- SEQ ID NOS:1-1079 were translated in all three reading frames, and the nucleotide sequences and translated amino acid sequences used as query sequences to search for homologous sequences in either the GenBank (nucleotide sequences) or Non-Redundant Protein (amino acid sequences) databases. Query and individual sequences were aligned using the BLAST 2.0 programs (National Center for Biotechnology Information, Bethesda, Maryland). (see also Altschul, et al. *Nucleic Acids Res.* (1997) *25*:3389-3402). The sequences were masked to various extents to prevent searching of repetitive sequences or poly-A sequences, using the XBLAST program for masking low complexity as described above in Example 1. --

Please replace the paragraph on page 48, lines 25-34, with the following rewritten paragraph:

A5

-- Some polynucleotides exhibited multiple profile hits where the query sequence contains overlapping profile regions, and/or where the sequence contains two different functional domains. Each of the profile hits of Table 3 are described in more detail below. The acronyms for the profiles (provided in parentheses) are those used to identify the profile in the Pfam and Prosite databases. The Pfam database can be accessed through web sites supported by the Washington University, St. Louis (Missouri), The Sanger Centre (United Kingdom); and The Karolinska Institute Center for Genomics Research. The Prosite database is publicaly available through the ExPASy Molecular Biology Server.

The public information available on the Pfam and Prosite databases regarding the various profiles, including but not limited to the activities, function, and consensus sequences of various proteins families and protein domains, is incorporated herein by reference. --

Please replace the paragraph on page 49, lines 25-35, with the following rewritten paragraph:

A6

-- ATPases Associated with Various Cellular Activities (ATPases; Pfam Accession No. PF0004). SEQ ID NOS:1035, 1058, and 1072 corresond to a sequence that encodes a member of a family of ATPases Associated with diverse cellular Activities (AAA). The AAA protein family is composed of a large number of ATPases that share a conserved region of about 220 amino acids containing an ATP-binding site (Froehlich *et al., J. Cell Biol.* (1991) *114*:443; Erdmann *et al. Cell*

(1991) *64*:499; Peters *et al., EMBO J.* (1990) *9*:1757; Kunau *et al., Biochimie* (1993) *75*:209-224; Confalonieri *et al., BioEssays* (1995) *17*:639; see also the AAA Server Homepage). The AAA domain, which can be present in one or two copies, acts as an ATP-dependent protein clamp (Confalonieri *et al.* (1995) *BioEssays 17*:639) and contains a highly conserved region located in the central part of the domain. The consensus pattern is: [LIVMT]-x-[LIVMT]-[LIVMF]-x-[GATMC]-[ST]-[NS]-x(4)-[LIVM]- D-x-A-[LIFA]-x-R. --

A6 condd

Please replace the paragraph on page 51, lines 19-33, with the following rewritten paragraph:

A7

-- Helicases conserved C-terminal domain (helicase C; Pfam Accession No. PF00271). SEQ ID NOS:227 and 1058 represent polynucleotides encoding novel members of the DEAD/H helicase family. The DEAD box family comprises a number of eukaryotic and prokaryotic proteins involved in ATP-dependent, nucleic-acid unwinding. All DEAD box family members of the above proteins share a number of conserved sequence motifs, some of which are specific to the DEAD family while others are shared by other ATP-binding proteins or by proteins belonging to the helicases 'superfamily' (Hodgman, *Nature* (1988) *333*:22 and *Nature* (1988) *333*:578. One of these motifs, called the 'D-E-A-D-box', represents a special version of the B motif of ATP-binding proteins. Some other proteins belong to a subfamily which have His instead of the second Asp and are thus said to be 'D-E-A-H-box' proteins (Wassarman D.A., et al., *Nature* (1991) *349*:463; Harosh I., et al., *Nucleic Acids Res.* (1991) *19*:6331; Koonin E.V., et al., *J. Gen. Virol.* (1992) *73*:989. The following signature patterns are used to identify member for both subfamilies: 1) [LIVMF](2)-D-E-A-D-[RKEN]-x-[LIVMFYGSTN]; and 2) [GSAH]-x-[LIVMF](3)-D-E-[ALIV]-H-[NECR]. --

Please replace the paragraph starting at page 55, line 32, with the following rewritten paragraph:

A8

-- WW/rsp5/WWP domain signature and profile (WW domain; Pfam Accession No. PF00397). SEQ ID NO:606 corresponds to a gene encoding a protein comprising a WW domain. The WW domain (Bork *et al. Trends Biochem. Sci.* (1994) *19*:531-533; Andre *et al. Biochem. Biophys. Res. Commun.* (1994) *205*:1201-1205; Hofmann *et al. FEBS Lett.* (1995) *358*:153-157; Sudol *et al. FEBS Lett.* (1995) *369*:67-71 (also known as rsp5 or WWP) was discovered as a short conserved region in a number of

unrelated proteins, among them dystrophin, the gene responsible for Duchenne muscular dystrophy. The domain, which spans about 35 residues, is repeated up to 4 times in some proteins. It has been shown (Chen *et al. Proc. Natl. Acad. Sci. U.S.A.* (1995) *92*:7819-7823) to bind proteins with particular proline-motifs, [AP]-P-P-[AP]-Y, and thus resembles somewhat SH3 domains. The WW domain conatins beta-strands grouped around four conserved aromatic positions, generally tryptophan. The name WW or WWP derives from the presence of two tryptophane as well as a conserved proline. The WW domain is frequently associated with other domains typical for --

A8 cmdd

Please replace Table 12 on page 65, lines 15-16, with the following revised table:

A9

--**Table 12:** Pools of Clones and Libraries Deposited with ATCC on or before September 23, 1999

Library No.	CMCC No.	ATCC Deposit No.	Library No.	CMCC No.	ATCC Deposit No.
ES55	5058	PTA-739	ES65	5068	PTA-749
ES56	5059	PTA-740	ES66	5069	PTA-750
ES57	5060	PTA-741	ES67	5070	PTA-751
ES58	5061	PTA-742	ES68	5071	PTA-752
ES59	5062	PTA-743	ES69	5072	PTA-753
ES60	5063	PTA-744	ES70	5073	PTA-754
ES61	5064	PTA-745	ES71	5074	PTA-755
ES62	5065	PTA-746	ES72	5075	PTA-756
ES63	5066	PTA-747	ES73	5076	PTA-757
ES64	5067	PTA-748	ES74	5077	PTA-758

--

IN THE CLAIMS:

Please cancel claims 1-12 without prejudice.

Please add the following new claims 13-102.

A10

--13. (New) An isolated polynucleotide comprising at least 15 contiguous nucleotides of a nucleotide sequence having at least 90% sequence identity to a sequence selected from the group consisting of: SEQ ID NO:137, a degenerate variant of SEQ ID NO:137, and a complement of SEQ ID NO:137.